

INTERNATIONAL JOURNAL OF RESEARCH & ENGINEERING

POSTAL APPLICATIONS

JOURNAL INTERNATIONAL DE RECHERCHE ET TECHNIQUE

APPLICATIONS POSTALES

INAUGURAL ISSUE

LINKING THE WORLD WITH TECHNOLOGY

ROBOTICS • VISION SYSTEMS • PATTERN RECOGNITION • NEURAL NETWORKS • COMPUTER ARCHITECTURE • INTELLIGENT SYSTEMS
ERGONOMICS • NETWORKING • ECONOMICS • ELECTRONIC COMMUNICATIONS • CODE PRINTING/READING • MANAGEMENT INFORMATION



ROBOTIQUES • SYSTÈMES DE VISION • RECONNAISSANCE DE FORMES • RÉSEAUX DE NEURONES • ARCHITECTURE DU CALCUL
SYSTÈME INTELLIGENT • TRAITEMENT DES OBJETS • ERGONOMIE • RESEAU • ECONOMIE • SYSTÈMES DE TRANSMISSION • IMPRESSION

TECHNOLOGIES DANS LE MONDE

PREMIER NUMERO

ACKNOWLEDGEMENTS

The Editors would like to recognize and express their appreciation to everyone involved in the publication of this journal. This includes USA Today, RICOH and San Francisco/Moscow Teleport for supporting an effort for an international transmission for the Journal; MIR Publications, USSR, for inputting the text of the papers into their publishing system; Radio i Svyaz, USSR, for their partnership effort with MIR Publications; and Hemisphere Publications for finalizing the journal information transmitted from the USSR. This team effort has allowed the inaugural issue of the IJRE to truly be an international publication.

Publication Editors

Dmitry Zaidin
MIR Publishers, USSR

Louis Fancher
Hemisphere Publishing Corporation
A Member of the Taylor & Francis Group

Terrence McCaffrey
U.S. Postal Service

Journal Coordinators

Kim Alcorn
Linda Mahoney
Evelena Carroll

The International Journal of Research and Engineering, Postal Applications (IJRE) (ISSN 1043-7134) is published semi-annually by the IJRE. For more information complete the enclosed information sheet or call (202) 268-3867. Copyright © 1989 by the IJRE. All rights reserved. Postmaster, please send address changes to the: IJRE; Office of Advanced Technology; 475 L'Enfant Plaza, SW; Washington, DC 20260-8118.

REMERCIEMENTS

Les rédacteurs en chef désirent exprimer leur reconnaissance à tous ceux qui ont participé à la publication de ce journal. Et parmi eux, à USA Today, à RICOH et à San Francisco/Moscou Teleport pour avoir collaboré à la transmission internationale pour le Journal ; aux Publications MIR, URSS pour avoir saisi le texte des articles dans leur système d'édition ; à Radio i Svyaz, URSS, pour l'aide qu'ils ont apportée aux publications MIR ; et aux Publications Hemisphere pour avoir mis au point les informations du journal en provenance de l'URSS. Ce travail d'équipe a permis au premier numéro de l'IJRE d'être une publication véritablement internationale.

Rédacteurs en chef des publications

Dmitry Zaidin
MIR Publishers, USSR

Louis Fancher
Hemisphere Publishing Corporation
A Member of the Taylor & Francis Group

Terrence McCaffrey
U.S. Postal Service

Coordinateurs du journal

Kim Alcorn
Linda Mahoney
Evelena Carroll

Le Journal International de Recherche et Technique, Applications Postales (IJRE) (ISSN 1043-7134) est une revue semestrielle publiée par l'IJRE. Pour obtenir des informations supplémentaires, veuillez compléter la carte de demande d'informations jointe ou téléphonez au (202) 268-3867. Droits de reproduction © 1989 par l'IJRE. Tous droits réservés. Monsieur le Receveur des Postes, ayez l'amabilité d'envoyer les changements d'adresse à : IJRE ; Office of Advanced Technology ; 475 L'Enfant Plaza, SW ; Washington, DC 20260-8118.

TABLE OF CONTENTS

Address Reading

Similarity-Invariant Analysis of Handwritten ZIP Code Using Fourier Descriptors <i>F. Ghorbel, G. Cazuguel and J.L. de Bougrenet de la Tocnaye</i>	1
Adaptation of Recognition Algorithms to Arbitrary Character Orientation <i>O.E. Plyatsek, L.E. Yaschook, Dr. Sci.</i>	13
Application of Neural-Net to Read Printed Japanese Addresses <i>Shinji Sase, Hiroyuki Kami, Toshio Ishikawa and Kanehiro Kubota</i>	25
A System to Locate and Recognize ZIP Codes in Handwritten Addresses <i>Sargur N. Srihari, Edward Cohen, Jonathan J. Hull and Leonard Kuan</i>	37
Towards Developing a Real-Time System to Locate Address Blocks on Mail Pieces <i>Sargur N. Srihari, Jung Soh and Paul W. Palumbo</i>	57
A Recognition Method of Printed Chinese Characters by Feature Combination <i>Zheng Zhang, Irmfried Hartmann, Jun Guo and Richard Suchenwirth</i>	77

Automatic and Adaptive Material Handling

Advanced Transport Stacker Live Mail Test <i>John M. Buday</i>	93
Microcomputer Motion Analysis Applied to Mail Transport Systems <i>L. Scot Duncan</i>	105
New Results of Research and Engineering Work on Pneumatic Conveying of Postal Parcels <i>Walter Merz</i>	117
Recovery of Mail Piece Shape from Range Images Using 3-D Deformable Models <i>Franc Solina and Ruzena Bajcsy</i>	125
Advanced Research on Techniques for Intelligent Singulation of Manual Letters <i>Ramojus P. Vaitys</i>	141

Expert Systems, Human Factors and Maintenance

A Human Factors Tool for Non-experts and its use in Site Planning for Datapost Inland Tracking and Tracing System <i>Jane Dillon and Joe Langford</i>	153
Effect of a Work Hardening Program in Mechanized Mail Sorting <i>Rose Oldfield Hayes, Ph.D.</i>	173
User Assistance Methods Within Expert Systems <i>Roger H. Whitton</i>	187

Management Information Systems, Economics and Planning

The Economic Balance between Window Clerk Workhours and Customer Waiting Time <i>J. Walter Lautenberger, Jr.</i>	199
---	-----

Abstracts in Other Languages

RECONNAISSANCE DE FORME D'OBJET A PARTIR D'UNE SERIE D'IMAGES PAR UTILISATION DE MODELES 3D DEFORMABLES

Franco Solina
Ruzena Bajcsy
GRASP Laboratory
Department of Computer and
Information Science
University of Pennsylvania
Philadelphia, PA 19104, U.S.A.
Téléphone (215) 898-6222
Téléfax (215) 898-1130

Résumé

Le tri automatisé de tous les types de courrier doit prendre en compte différentes formes et tailles d'objets. On propose un système de classement des objets en fonction de leur forme en se servant d'une série d'images.

Des modèles volumétriques de lettres seules caractérisés par des paramètres de position, orientation, taille et forme, sont reconnus en utilisant une minimisation par moindres carrés d'une fonction de correspondance de modèle. Les modèles utilisés sont des superquadrics avec des déformations globales. Les paramètres reconnus peuvent servir aux différents schémas de classement dont dépend la méthode particulière utilisée pour la manipulation du flux de courrier. Un schéma de classement qui sépare les lettres en objets plats, roulés, parallélépipédiques et lettres irrégulières, est présenté.

1. Introduction

Le tri automatisé de toutes les lettres est un problème difficile parce que les lettres diffèrent largement en taille et en forme. En général, on doit connaître la localisation, l'orientation, la taille et la forme d'une lettre pour commencer la procédure correcte de saisie.

La vision par ordinateur en tant que méthode pour la localisation et la description d'objets sans contact physique direct semble être la bonne approche pour faire cela d'une manière rapide et sûre.

On s'est servi avec succès de la vision par ordinateur dans beaucoup d'applications industrielles. Les méthodes, utilisées dans la majorité de ces systèmes industriels de vision, ne peuvent cependant être appliquées au problème de classification des lettres. La plupart de ceux-ci appelés systèmes de reconnaissance d'objet utilisant des modèles (« model-based »)

comptent sur un jeu de modèles précis et immuables de tous les objets que le système est capable de reconnaître. Sur la base de quelques caractéristiques prédéfinies reconnues dans les images, on sélectionne des hypothèses dont les modèles sont projetés sur l'image, pour vérifier s'ils correspondent avec le reste des caractéristiques de l'image (pour un exemple d'une telle méthode [Bolle et Horaud, 1986]). Cette approche est possible uniquement dans des environnements étroitement contrôlés où la forme des objets est bien définie à l'avance et/ou le nombre des différents objets est petit. Cela n'est évidemment pas le cas des lettres qui sont de taille et de forme variées. Quand des modèles précis d'objets ne sont pas disponibles, la vision par ordinateur traditionnelle préconise une prudente réduction de donnée (Marr, 1982). Tout d'abord des modèles de formes élémentaires tels les côtes, les coins et des éléments de surface sont analysés localement.

A cause de la faible granularité de ces modèles, un grand nombre d'entre eux est nécessaire même pour des descriptions de scènes simplistes. Pour faciliter tout raisonnement concernant la scène, ces modèles locaux doivent être intégrés dans des entités plus larges, des cylindres généralisés étant le modèle le plus courant, qui correspondent à des parties d'objets ou des objets. Dans le cas des lettres, cette fusion des modèles locaux est difficile et sujette à erreur parce que les lettres, tout comme les scènes naturelles, ne se conforment pas à des formes parfaitement géométriques à cause des bords arrondis, des coins tordus, des côtes bombés et des surfaces rugueuses. Il est difficile d'extraire avec sûreté le modèle grande échelle le plus approprié à partir du grand nombre de modèles élémentaires entachés de défauts. Nous croyons qu'une telle approche ascendante pour classer les lettres à partir des images en fonction de leur forme ne réussirait pas.

Lors de la deuxième conférence du Service Postal des Etats-Unis sur les technologies avancées (Solina et Bajcsy, 1986), nous avons proposé d'appliquer directement des modèles volumétriques de plus grande granularité — des modèles superquadrics avec des déformations — pour l'interprétation des lettres. L'information supplémentaire nécessaire pour pallier l'information manquante et rejeter l'information locale erronée peut être fournie par des contraintes implicites des modèles volumétriques compacts pourvu qu'ils aient la bonne granularité. Les superquadrics fournissent un vocabulaire de formes bien adapté à la description des formes du monde réel au niveau de détail qui correspond à la notion humaine de parties (Pentland, 1986). Un modèle superquadric est en général suffisant pour modéliser une lettre. Les paramètres de ces modèles

superquadrics peuvent être trouvés directement à partir des images (Solina, 1987). La reconnaissance peut être intuitivement expliquée en termes de forces extrinsèques et intrinsèques. Les forces extrinsèques adaptent les modèles dont les formes possibles et l'arrangement des paramètres du modèle sont gouvernés par les forces internes. Puisque les superquadrics peuvent modéliser la plupart des formes géométriques courantes, tels les parallélépipèdes, les cylindres, les ellipses et les formes intermédiaires, ils sont des modèles appropriés pour la plupart des lettres.

La suite du rapport est organisée comme suit : le paragraphe 2 est une brève introduction aux superquadrics, le paragraphe 3 décrit la reconnaissance du modèle à partir des images, et le paragraphe 4 traite du classement. La discussion dans le paragraphe 5 compare les avantages et les défauts de la méthode et introduit des extensions futures, en particulier concernant l'utilisation d'une telle reconnaissance de forme pour la segmentation. Nous avons utilisé des images provenant de différentes sources. Cependant toutes les images utilisées pour cette étude furent obtenues à l'aide d'un dispositif laser d'acquisition d'image construit à l'université de Pennsylvanie (Tsikos, 1987).

2. Superquadrics

Les superquadrics sont une extension des surfaces quadrics de base et des solides. Les superquadrics ont été considérés comme les primitives pour la représentation des formes dans le graphisme informatisé (Barr, 1981) et la vision par ordinateur (Pentland, 1986, Bajcsy et Solina, 1987, Boulton et Gross, 1987). Les superquadrics peuvent être comparés à des mottes d'argiles que l'on peut en plus déformer et coller ensemble pour obtenir des modèles réalistes comme cela a été joliment démontré sur le système graphique Supersketch de Pentland (Pentland, 1986).

La surface superquadric est définie par l'équation suivante :

$$F(x, y, z) = \left(\left(\left(\frac{x}{a_1} \right)^{\frac{2}{\epsilon_1}} + \left(\frac{y}{a_2} \right)^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_1}{\epsilon_1 + \epsilon_2}} + \left(\frac{z}{a_3} \right)^{\frac{2}{\epsilon_1}} \right)^{\epsilon_1} \quad (1)$$

quand ϵ_1 et ϵ_2 valent 1, la surface définie est une ellipsoïde ou, si a_1, a_2, a_3 sont égaux, une sphère. Quand $\epsilon_1 \ll 1$ et $\epsilon_2 = 1$, la surface superquadric a la forme d'un cylindre. On obtient des parallélépipèdes quand $\epsilon_1 \ll 1$ et $\epsilon_2 \ll 1$. Les possibilités de modelage des superquadrics peuvent être mises en évidence en les déformant de différentes manières, comme les tailler en pointe et les courber (Solina, 1987). Quelques exemples de modèles superquadrics apparaissent sur la figure 1. La fonction mise en équation (1) est appelée fonction intérieur-extérieur parce qu'elle détermine où un point donné (x, y, z) se trouve par rapport à la surface superquadric. Si $F(x, y, z) = 1$, le point (x, y, z) se trouve sur la surface du superquadric.

Si $F(x, y, z) > 1$, le point correspondant se trouve à l'extérieur et si $F(x, y, z) < 1$, le point correspondant se trouve à l'intérieur du superquadric.

La fonction intérieur-extérieur (1) définit la surface superquadric pour un système de coordonnées centré sur l'objet (x_s, y_s, z_s). D'un autre côté, les points 3D dans les images sont exprimés dans un système de coordonnées de l'image. Pour retrouver un superquadric dans une position quelconque, une fonction intérieur-extérieur pour position quelconque est utilisée dans laquelle la relation entre le système de coordonnées

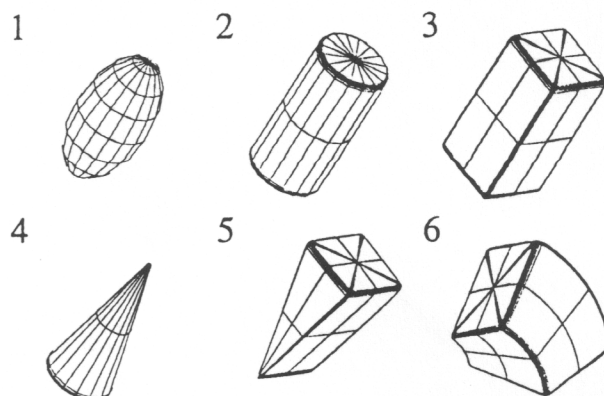


FIGURE 1

de l'image et le système de coordonnées centré sur l'objet est décrite par une transformation homogène T . Nous exprimons les éléments de la rotation de la matrice de transformation T avec les angles d'Euler (ϕ, θ, ψ) (Paul, 1981). La fonction intérieur-extérieur pour les superquadrics en position quelconque est alors

$$F(x, y, z) = F(x, y, z; a_1, a_2, a_3, \epsilon_1, \epsilon_2, \phi, \theta, \psi, p_x, p_y, p_z). \quad (2)$$

Cette fonction intérieur-extérieur étendue a 11 paramètres ; a_1, a_2, a_3 définissent la taille du superquadric ; ϵ_1 et ϵ_2 sont des paramètres de formes ; ϕ, θ, ψ définissent l'orientation dans l'espace, et p_x, p_y, p_z définissent la position dans l'espace. Nous appelons l'ensemble de tous les paramètres du modèle : $\Lambda = \{a_1, a_2, \dots, a_{11}\}$.

3. Reconnaissance de superquadrics

Pour la reconnaissance de forme des lettres nous faisons l'hypothèse que juste une seule lettre est présente dans le champ de l'image à un instant donné. Une telle séparation des lettres peut être obtenue par des moyens mécaniques. Tous les points qui représentent la surface d'appui dans l'image peuvent être extraits de l'ensemble de tous les points en faisant correspondre un plan à la surface d'appui et en retirant tous les points sur ce plan ou proche du plan. Nous pouvons supposer que tous les points restants appartiennent à la surface de la lettre. Supposons que nous avons N points 3D de surface (x_w, y_w, z_w) que nous souhaitons modéliser par un superquadric. Nous souhaitons faire varier les 11 paramètres a_j , $j = 1, \dots, 11$, dans l'équation 2, pour obtenir des valeurs des différents paramètres a_j telles que la plupart des points 3D se situeraient sur la surface du modèle ou proche de celle-ci.

Il n'y aura probablement pas un ensemble de paramètres 1 qui correspondent parfaitement aux données. Trouver le modèle 1, pour lequel la distance des points à la surface du modèle soit minimale, est un problème de minimisation par moindres carrés. A cause du problème de l'auto-occlusion, tous les côtés d'un objet ne sont pas visibles en même temps, nous devons introduire une contrainte supplémentaire. Parmi toutes les solutions possibles nous désirons trouver le superquadric le plus petit qui corresponde à tous les points donnés au sens moindres carrés du terme. Nous définissons la fonction

suivante qui possède un minimum correspondant au plus petit superquadric qui corresponde à l'ensemble des points 3D et une valeur de la fonction pour les points de la surface, comme avant minimisation.

$$R = \sqrt{a_1 a_2 a_3} (F - 1). \quad (3)$$

Puisque pour un point (x_w, y_w, z_w) sur la surface d'un superquadric

$$R(x_w, y_w, z_w; a_1, \dots, a_{11}) = 0. \quad (4)$$

Nous devons trouver

$$G = \min \sum_{i=1}^N [R(x_{w_i}, y_{w_i}, z_{w_i}; a_1, \dots, a_{11})]^2. \quad (5)$$

Puisque R est une fonction non linéaire de 11 paramètres $a_j, j = 1, \dots, 11$, la minimisation doit se faire de manière itérative. Ayant établi un jeu d'essai des valeurs des paramètres Λ_k du modèle, nous évaluons l'équation (3) pour tous les N points et employons une procédure pour améliorer la solution d'essai. La procédure est alors répétée avec un jeu de nouvelles valeurs d'essai Λ_{k+1} jusqu'à ce que la somme des moindres carrés cesse de décroître ou que les changements ne soient plus statistiquement significatifs. Dans la plupart des cas 15 itérations sont plus que suffisantes. Nous utilisons la méthode de Levenberg-Marquardt pour la minimisation non linéaire par moindres carrés (Press et al, 1986), puisque les premières dérivées $\delta R / \delta a_i$ ai pour $i = 1, \dots, 11$ peuvent être calculées analytiquement.

Il suffit d'une première estimation très approximative sur la position réelle de l'objet, son orientation et sa taille pour assurer la convergence vers un minimum local qui correspond à la forme réelle. Ceci est important puisque ces paramètres ne peuvent être estimés qu'à partir de l'ensemble des points appartenant au côté visible de l'objet et par conséquent on ne peut démarrer avec des estimations très précises. Les valeurs initiales pour les deux paramètres de forme, ϵ_1 et ϵ_2 , peuvent toujours être égales à 1, ce qui signifie que le modèle initial Λ_E est toujours une ellipsoïde. La position dans les coordonnées universelles est estimée en calculant le centre de gravité de tous les points, et l'orientation est estimée en calculant les moments centraux en tenant compte du centre de gravité. Nous orientons le modèle initial Λ_E de sorte que l'axe z du système de coordonnées, centré sur l'objet, s'aligne avec le plus grand côté (axe de moindre inertie) de l'objet. Ceci parce que déforme les objets, en les courbant et en les taillant en pointe, les affecte normalement le long de leur plus grand côté. Les estimations concernant la taille du modèle sont simplement fournies par l'étendue des points le long du nouvel axe de coordonnées. Les figures 2, 3, 4 et 5 donnent des exemples de reconnaissance de modèle pour chacun des quatre types de courrier proposés : une boîte, un rouleau, une lettre plate et un colis irrégulier. La faible correspondance du premier modèle sans déformation de la figure 5 peut être améliorée en appliquant des déformations globales aux modèles superquadrics.

On peut reconnaître des superquadrics déformés en utilisant la même technique que pour la reconnaissance de superquadrics non déformés. La seule différence réside dans le fait que quelques paramètres supplémentaires décrivant la déformation doivent aussi être retrouvés.

Des déformations simples telles qu'un taillage en pointe, une courbure et un vrillage ne nécessitent juste que peu de paramètres supplémentaires (Barr, 1984). La déformation est une fonction D

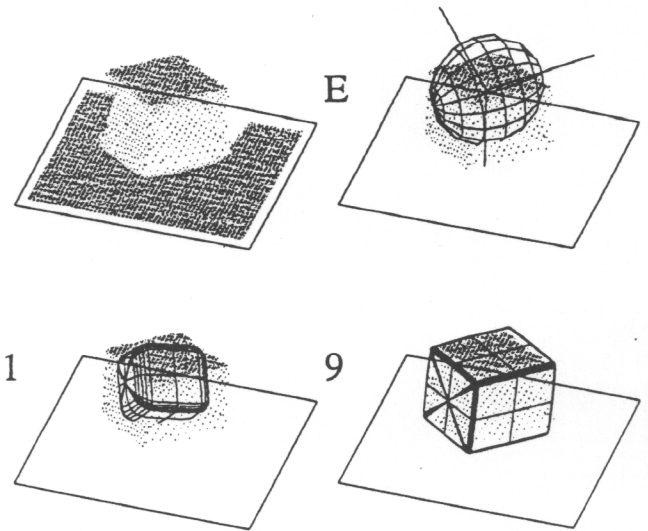


FIGURE 2

qui modifie explicitement l'ensemble des coordonnées des points dans l'espace :

$$\mathbf{X} = D(\mathbf{x}), \quad (6)$$

où \mathbf{x} représente les points du solide avant déformation et \mathbf{X} les points correspondants après déformation. \mathbf{x} et \mathbf{X} sont exprimés dans le système de coordonnées centré sur l'objet. Toute translation ou rotation est effectuée après la déformation. Un modèle taillé en pointe et courbé peut être décrit schématiquement comme suit :

$$\text{Trans}(\text{Rot}(\text{Band}(\text{Tapers}(\mathbf{x}))))). \quad (7)$$

La structure de ce modèle semble avoir une signification perceptible. Quand on construit un modèle, on doit procéder de l'intérieur vers l'extérieur, en prenant tout d'abord un modèle sans déformation, en le remodelant et en le mettant dans la position absolue correcte pour toutes les coordonnées. Quand on cherche

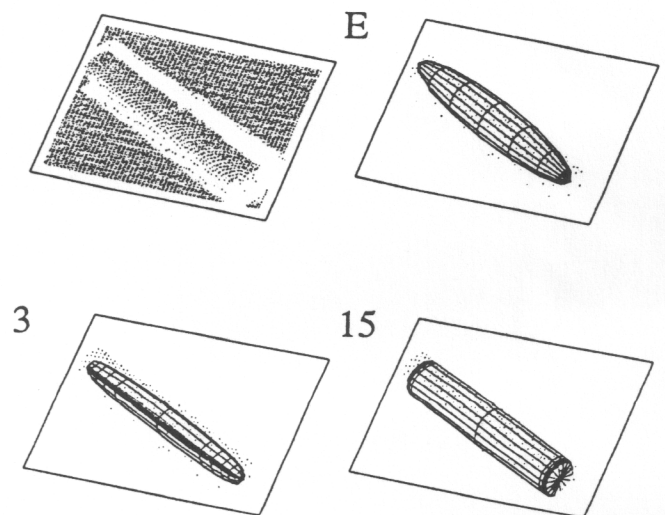


FIGURE 3

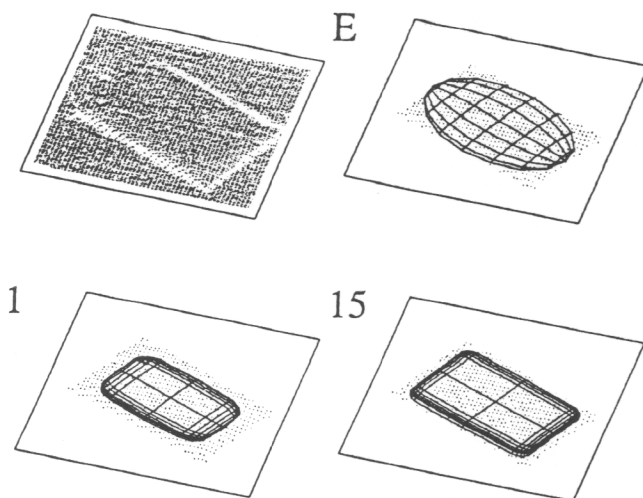


FIGURE 4

à reconnaître un modèle, l'opération doit s'effectuer en sens inverse. On doit tout d'abord repérer la position et l'orientation de l'objet modélisé, dans l'espace, avant de reconnaître sa forme. La fonction intérieur-extérieur du modèle déformable que nous utilisons a 4 paramètres supplémentaires : deux paramètres de déformation en pointe, l'un pour la déformation le long de l'axe z , l'autre indépendant pour la déformation dans la direction des axes x et y , et deux paramètres de courbure : α , un angle qui définit l'orientation du plan de courbure autour de l'axe z , et β , l'angle de courbure effectif :

$$F(x, y, z) = F(X, Y, Z; a_1, \dots, a_{11}, K_x, K_y, \alpha, \beta). \quad (8)$$

La fonction de correspondance (3) peut être considérée comme une fonction énergie dans l'espace des paramètres de modèle. Les méthodes de minimisation ne peuvent, en général, seulement garantir qu'une convergence vers un minimum local. La position de départ dans l'espace des paramètres (Λ_E) déter-

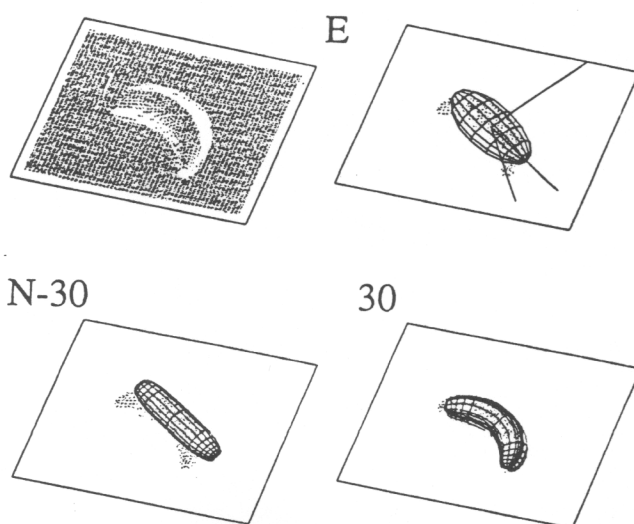


FIGURE 5

mine vers quel minimum la procédure de minimisation convergera. Nous devons nous assurer que la procédure de minimisation ne s'arrête pas prématurément sur un minimum local peu profond, mais trouve plutôt un minimum significatif. On évite les solutions correspondant à des minimums locaux peu profonds en ajoutant du bruit de fond à la valeur de la fonction de correspondance du modèle accepté à chaque itération durant la reconnaissance du modèle. Quand on compare la valeur contaminée par le bruit avec le degré d'aptitude du modèle pour l'itération en cours, un nouveau modèle peut être sélectionné bien qu'il ne corresponde pas aux données, aussi bien que l'ancien modèle. Cette technique stochastique d'insertion de « jitter » dans la procédure de correspondance ressemble à un recuit simulé (voir figure 7). La partie dans la reconnaissance de modèle, qui nécessite le plus de temps, est l'évaluation de la fonction de correspondance et de tous ses dérivés partiels pour chacun des points pendant chaque itération. Quand la somme des moindres carrés est une fonction qui augmente de manière monotone, cela achève le contrôle de la somme partielle après chaque addition. Dès que la somme devient plus grande que la somme des moindres carrés du modèle accepté et contaminé par le bruit, cela n'a plus de sens de continuer. Le modèle ne peut être accepté.

On peut gagner substantiellement en vitesse en sous-échantillonnant l'image d'origine et en utilisant une série de grilles grossières puis fines durant la minimisation.

Les modèles reconnus à partir d'images à grille plus lâche peuvent encore être une bonne représentation de l'objet (figure 6). Durant la reconnaissance itérative du modèle, la fonction de correspondance chute typiquement très rapidement jusqu'à atteindre un plateau. De nouvelles itérations n'apportent pas d'améliorations substantielles dans le degré de correspondance (figure 7). On peut mener un calcul rapide et efficace sur une hiérarchie de grilles grossières. Nous avons réalisé un mécanisme de reconnaissance de modèle multi-résolution qui démarre sur une image très grossière. Lorsque l'on n'améliore plus le degré de correspondances, la minimisation continue sur une image à plus haute densité jusqu'à ce que l'on atteigne le maximum de résolution ou l'image d'origine (figure 7). La méthode multi-résolution est plus rapide parce qu'elle nécessite moins de temps de calcul à chaque itération. Le nombre d'itérations peut, cependant, ne pas être plus petit. Le nombre d'itérations peut même être plus grand parce que, durant la reconnaissance multi-résolution, les modèles pour des images plus clairsemées convergent vers des ensembles de paramètres un peu différents du jeu de paramètres résultant sur la grille la plus fine. Une réalisation de la procédure de reconnaissance sur une architecture entièrement parallèle serait appropriée, puisque l'évaluation de la fonction de correspondance et de ses dérivés partiels est localement indépendante. La reconnaissance des modèles présentés dans ce rapport, dont le nombre total de points pour chacun d'eux est de l'ordre de plusieurs centaines, prend à peu près 20 secondes de temps CPU sur un calculateur VAX 785.

Nous avons testé la fiabilité de la méthode de reconnaissance en prenant une série d'images du même objet dans des positions et orientations différentes. Les modèles reconnus sont très favorablement comparables. Pour les détails et résultats concernant la fiabilité, la stabilité et l'ambiguïté de la reconnaissance de modèle voir (Solina, 1987).

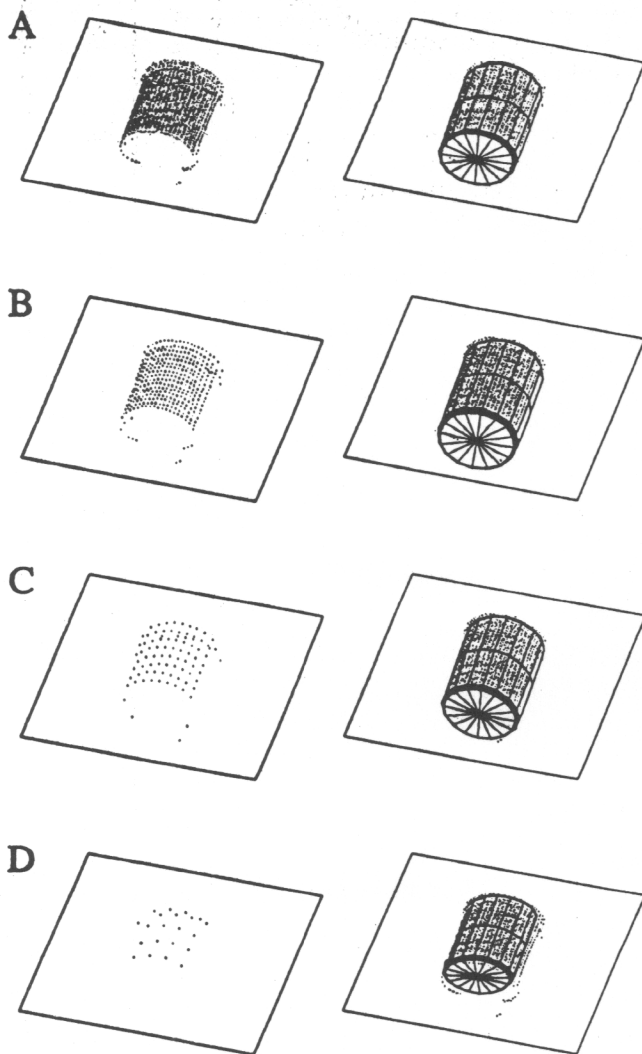


FIGURE 6

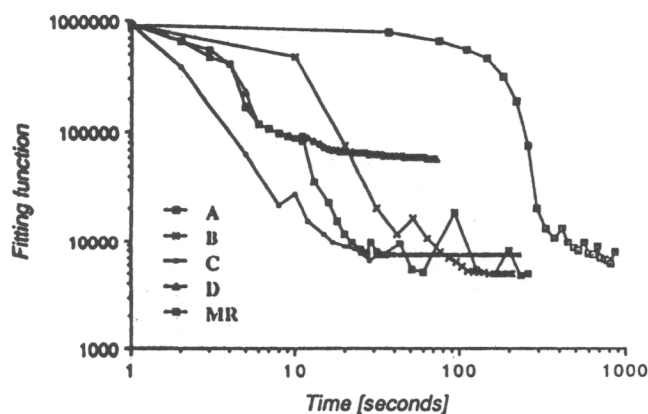


FIGURE 7

4. Classification

Les résultats de la procédure de reconnaissance de modèle sont la position et l'orientation des modèles, ainsi que leurs paramètres de taille et de forme. L'espace des paramètres est un espace continu, mais quelques groupes de paramètres correspondent à des primitives géométriques communes telles des parallélépipèdes et des cylindres. Pour ces classes ou catégories, les différences de paramètres à l'intérieur d'une catégorie paraissent plus petites que les différences de paramètre entre catégories même quand elles sont de la même taille. En faisant la cartographie des symboles sur l'espace continu des paramètres superquadrics, il est possible de définir des classes distinctes d'objets. Pour la manipulation manuelle des lettres, on distingue quatre types de lettres, reflétant les changements que l'on rencontre habituellement dans la structure des formes de lettres.

Les quatre classes sont les lettres plates, les boîtes, les tubes et les lettres irrégulières qui forment une classe composée des éléments qui n'appartiennent à aucune des trois premières classes. Nous pourrions définir plusieurs schémas différents de classification basés sur les paramètres reconnus.

Bien que la nature des équipements de manipulation automatique du courrier doit être prise en compte, il semble raisonnable de prendre en considération la classification mentionnée précédemment, puisqu'elle reflète déjà quelques critères sensibles pour le matériel de manipulation.

Nous pouvons faire davantage que simplement classer les lettres en différents groupes en fonction de leur forme. A partir des deux paramètres de forme et de taille du modèle, on peut calculer le rayon de courbure exact sur les côtés, pour évaluer la netteté ou la rondeur des côtés-paramètres qui sont importants pour la manipulation automatique. La méthode de minimisation par moindres carrés pour la reconnaissance de modèle nous fournit le niveau de qualité du modèle représentant l'objet réel. Ce reste G provenant de l'équation (5), après réduction par moindres carrés, mesure la qualité de la correspondance qui joue un rôle important dans la classification.

Le modèle reconnu peut avoir la forme d'un parallélépipède mais le niveau de qualité de la correspondance peut être très faible, ce qui indique que l'objet réel est irrégulier (un paquet enveloppé, par exemple).

Le modèle reconnu dans ce cas est seulement une approximation grossière de la forme réelle mais suffisante, par exemple, pour saisir l'objet. Une faible correspondance du modèle, qui provoque le classement d'un objet en tant que paquet irrégulier, pourrait aussi indiquer que l'on aurait intérêt à appliquer des déformations globales pour améliorer la description de la forme. Des lettres, qui sont normalement classées ensemble dans la classe des lettres irrégulières, peuvent être mieux décrites si on déforme, en les taillant en pointe ou en les courbant, les modèles superquadrics, comme cela nous est montré en la figure 5.

En se basant sur la taille du modèle sur les trois axes (a_0 , a_1 , a_2), sur les deux paramètres de forme (ϵ_1 , ϵ_2) et sur la qualité de la correspondance G , nous établissons les règles de classification dans le Tableau 1.

TABLE 1

input $a_1, a_2, a_3, \epsilon_1, \epsilon_2, G,$ $T_{FLAT}, W_{FLAT}, S_{BOX}, D_{ROLL}, L_{ROLL}, RES$
if $G > RES$ then <i>mail piece is an IRREGULAR MAIL PIECE</i>
else if $((a_1 < T_{FLAT} \text{ and } a_2, a_3 > W_{FLAT}) \text{ or } (a_2 < T_{FLAT} \text{ and } a_1, a_3 > W_{FLAT}))$ then <i>mail piece is a FLAT</i>
else if $[a_1 > S_{BOX} \text{ and } a_2 > S_{BOX} \text{ and } a_3 > S_{BOX} \text{ and } \epsilon_1 < 0.5 \text{ and } \epsilon_2 < 0.5]$ then <i>mail piece is a BOX</i>
else if $[a_1 > D_{ROLL} \text{ and } a_2 > D_{ROLL} \text{ and } a_3 > L_{ROLL} \text{ and } \epsilon_1 < 0.5 \text{ and } \epsilon_2 > 0.5]$ then <i>mail piece is a ROLL</i>
else <i>mail piece is an IRREGULAR MAIL PIECE</i>

5. Discussion

Le vocabulaire de forme proposée est destiné à des descriptions grossières d'objets, adaptées à la classification des lettres par leur forme. Les objets dont la partie cachée n'est pas symétrique de la partie visible peuvent ne pas être représentés convenablement. Bien que les déformations de base : tailler en pointe et courber, sont souvent suffisantes, elles ne couvrent pas tous les cas possibles. Un plus grand nombre de différentes déformations pourrait être utilisé, mais cela nécessiterait un plus grand nombre de paramètres de déformation. Le modèle que nous utilisons semble être un bon compromis, pas trop compliqué pour la reconnaissance, et convenable pour saisir et manipuler les objets.

Une densité de données non uniforme et un grand nombre de prises de vue très particulières dans les images furent un autre problème auquel nous avons dû faire face. Une densité de points non uniforme fait que les parties à plus haute densité influencent davantage la forme du modèle retenu que les parties à plus basse densité. Cependant la méthode de reconnaissance de modèle est tout à fait robuste à cet égard. Notons qu'aucune donnée, appartenant aux parties cachées, n'est disponible pour être utilisée au départ. La prudence s'impose encore. Quand on présente au système de reconnaissance de forme une vue particulière qui apparaît, par exemple, quand on aperçoit seulement une face d'un cube à partir des points présents sur l'image, on reconnaît alors un parallélépipède très mince qui correspond à cette face. Malheureusement, des images prises avec un système d'acquisition passif, qui utilise la triangulation, comportent plus de vues particulières que les images normales qui utilisent l'intensité lumineuse. Plus la distance (angle) entre la source d'éclairage et la caméra du système d'acquisition d'image est grande, meilleure est la précision des mesures, mais plus la vue ressemble à celle d'un objet particulier, puisque pour acquérir tous les points d'une partie de surface, cette partie doit être éclairée par le laser et vue par

la caméra en même temps. Une manière d'éliminer les vues particulières est de combiner plusieurs images. Habituellement, nous résolvons le problème des vues particulières en projetant les points visibles sur la surface support et nous les utilisons ensemble avec le reste de tous les points pour leur faire correspondre un modèle. Ceci résout le problème du manque de points dû à la vue particulière mais peut déformer quelques objets, tels des cylindres posés sur leur côté (la face cachée d'un cylindre est plate au lieu d'être cylindrique). On résout habituellement le problème en trouvant un modèle pour les deux cas ; tout d'abord nous faisons correspondre un modèle seulement à l'ensemble des points visibles, et puis de nouveau à l'ensemble des points visibles et de leurs projections sur la surface support. Si la qualité de la correspondance est à peu près la même dans les deux cas, le modèle avec le plus gros volume est choisi. Sinon, on prend le modèle avec la meilleure qualité de correspondance. Une autre manière de résoudre le problème des vues particulières est de prendre en compte la structure de la scène environnante. Les objets restent sur des supports, ils peuvent se toucher mais normalement ils ne s'interpénètrent pas. Witkin, Fleicher et Barr ont développé une méthode élégante pour décrire les contraintes géométriques entre les parties d'objets, ou objets, en termes d'énergie (Witkin et al., 1987). Si ce paradigme était utilisé pour la reconnaissance de forme des lettres, on devrait minimiser une somme de termes d'énergie, l'un d'eux étant la fonction de correspondance pour le modèle de l'objet, alors que les autres termes définiraient les contraintes sur les différentes relations géométriques, telles que le modèle de l'objet doit toucher la surface support en trois points au moins non alignés.

Dans cette étude nous nous sommes intéressés uniquement à la reconnaissance de forme de lettres isolées. Quand plusieurs lettres superposées peuvent être présentes dans la scène, la scène doit être segmentée. Chaque lettre serait représentée par un modèle unique. La segmentation dépend cependant de la forme de chacune des parties. Mais pour reconnaître la forme des parties, on doit connaître les points qui appartiennent à la même partie. Dès lors segmentation et reconnaissance de forme sont intimement liées. A cause de cette interdépendance, nous croyons que la segmentation et la reconnaissance de forme de chacune des parties devraient être faites simultanément. La segmentation, en tant que partie intégrante d'une étape de la reconnaissance de forme, est possible (Solina, 1987, Pentland, 1987) mais doit être faite de manière plus sûre et robuste avant que des applications pratiques puissent être envisagées.

Remerciements

Ce travail était soutenu par le contrat du Service Postal des Etats-Unis 104230-87-H-001/M-0195 et la subvention Darpa N00014-88-K-0630.

Références

- BAJCSY, R. AND SOLINA, F. 1987. « Three dimensional shape representation revisited », in *Proceedings First International Computer Vision Conference*, London, England, 231-241.
- BARR, A. H. 1981. « Superquadrics and angle-preserving transformations ». *IEEE Computer Graphics and Applications*, vol. 1, pp. 11-23.

BARR, A. H. 1984. « Global and local deformations of solid primitives ». *Computer Graphics*, vol. 18, no. 3, pp. 21-30.

BOLLES, R. C. AND HORAUD, P. 1986. « 3DPO : A three-dimensional part orientation system ». *International Journal of Robotics Research*, vol. 5, no. 3, pp. 3-26.

BOULT, T. E. AND GROSS, A. D. 1987. « Recovery of superquadrics from depth information », in *Proceedings Spatial Reasoning and Multi-Sensor Fusion Workshop*, St. Charles, IL, pp. 128-137.

MARR, D. 1982. *Vision* (Freeman, San Francisco).

PAUL, R. 1981. *Robot manipulators* (MIT Press, Cambridge, MA).

PENTLAND, A. P. 1986. « Perceptual organization and the representation of natural form ». *Artificial Intelligence*, vol. 28, no. 3, pp. 293-331.

PENTLAND, A. P. 1987. « Recognition by parts », in *Proceedings First International Computer Vision Conference*, London, England, pp. 612-620.

PRESS, W. H. et al. 1986. *Numerical recipes* (Cambridge University Press, Cambridge).

SOLINA, F. AND BAJCSY, R. 1986. « Modelling of mail pieces with superquadrics », in *Proceedings Second USPS Advanced Technology Conference*, Washington, DC, pp. 472-481.

SOLINA, F. 1987. « Shape recovery and segmentation with deformable part models ». Ph.D. dissertation, University of Pennsylvania, Philadelphia, USA.

TSIKOS, G. 1987. « Segmentation of 3-d scenes using multi-modal interaction between machine vision and programmable mechanical scene manipulation ». Ph.D. dissertation, University of Pennsylvania, Philadelphia, USA.

WITKIN, A. P. et al. 1987. « Energy constraints on parameterized models », in *Proceedings SIGGRAPH-87*, Anaheim, CA.

Franco Solina



Franco Solina a obtenu son Dipl.Ing. (diplôme d'ingénieur) et son M.S. (maîtrise ès sciences) d'ingénieur électricien à l'Université de Ljubljana, Yougoslavie en 1979 et 1982, respectivement. Il a obtenu son Ph. D. (doctorat) d'informatique à l'Université de Pennsylvanie, Philadelphie, U.S.A., en 1987 et est maintenant assistant d'informatique à l'Université de Ljubljana. Ses intérêts en matière de recherche résident dans les modèles pour la représentation de formes et la segmentation de la vision. Franco Solina est membre de l'Association américaine d'intelligence artificielle.

Ruzena Bajcsy



Ruzena Bajcsy a obtenu son premier Ph.D. (doctorat) d'ingénieur électricien à l'Université technique slovaque de Bratislava, Tchécoslovaquie en 1967 et son deuxième Ph.D. (doctorat) en informatique à l'Université de Stanford en 1972. Elle est professeur d'informatique à l'Université de Pennsylvanie où elle dirige le laboratoire « GRASP » et, actuellement, elle est également la directrice du Département d'informatique. Ses principaux intérêts en matière de recherche sont la perception active, l'intégration sensorielle et la segmentation visuelle. Ruzena Bajcsy est éditrice-adjointe de « Transactions IEEE sur l'analyse de modèles-type », « Intelligence de la machine », « Lettres de reconnaissance de modèles », « Journal de robotique et de vision des ordinateurs », « Journal du traitement de l'image et des graphiques ».

Tableau 1 : Règle de classification pour séparer les lettres en fonction de leur forme dans les quatre classes suivantes : boîtes, rouleaux, lettres plates et lettres irrégulières. Les paramètres d'entrée sont les paramètres superquadrics trouvés $a_1, a_2, a_3, \epsilon_1, \epsilon_2$ et le reste de la correspondance par moindre carré G . Une lettre est irrégulière soit si la qualité du niveau de correspondance G du modèle trouvé n'est pas suffisante, représenté par la constante RES , ou si les dimensions n'entrent pas dans les limites de taille fixées. La constante T_{FLAT} représente l'épaisseur maximale d'une lettre plate. La constante W_{FLAT} représente la longueur et la largeur minimales d'une lettre plate. La constante S_{BOX} représente la taille maximale de n'importe quelle dimension pour une boîte. La constante D_{ROLL} représente le diamètre minimum d'un rouleau et la constante L_{ROLL} représente la longueur minimale d'un rouleau. On a pris en compte uniquement, dans ce tableau de classification, les bornes inférieures des dimensions pour les différentes classes de lettres, bien que l'on aurait pu introduire aussi les bornes supérieures, qui dépendent soit des réglages ou des capacités de l'équipement qui saisit le courrier.

Figure 1 : Superquadrics ($1 - \epsilon_1 = \epsilon_2 = 1$; $2 - \epsilon_1 = 0,1, \epsilon_2 = 1$; $3 - \epsilon_1 = \epsilon_2 = 0,1$) et superquadrics déformés (les modèles 4 et 5 sont taillés en pointe et le modèle 6 est courbé).

Figure 2 : Reconnaissance de forme d'un objet ressemblant à un parallélépipède (une boîte). En haut on trouve l'image d'origine suivie par la séquence de reconnaissance montrant la première estimation (E) et les modèles après la première et la neuvième itération durant laquelle tous les 11 paramètres du modèle furent affinés. La séquence de reconnaissance du modèle ci-dessus prend à peu près seulement 20 secondes sur un ordinateur VAX 785. Les paramètres trouvés furent : $a_1 = 54$ mm, $a_2 = 55$ mm, $a_3 = 63$ mm, $\epsilon_1 = 0,1, \epsilon_2 = 0,1$. En accord avec les règles de décision du Tableau 1, cet objet fut classé comme une boîte de largeur = 126 mm, profondeur = 108 mm et hauteur = 110 mm.

Figure 3 : Reconnaissance de forme d'un journal roulé en forme de tube. En haut on trouve l'image d'origine suivie par la séquence de reconnaissance montrant la première estimation (E) et les modèles après la 3^{ème} et la 15^{ème} itération quand

tous les 11 paramètres du modèle furent ajustés simultanément. Les paramètres trouvés furent : $a_1 = 24$ mm, $a_2 = 28$ mm, $a_3 = 137$ mm, $\epsilon_1 = 0,1$, $\epsilon_2 = 1,1$.

En accord avec les règles de décision du Tableau 1, cet objet fut classé comme un rouleau de *longueur* = 274 mm, et *diamètre* = 52 mm.

Figure 4 : Reconnaissance de forme d'une lettre plate. En haut on trouve l'image d'origine suivie par la séquence de reconnaissance de modèle montrant l'estimation initiale du modèle (E) et les modèles après la 1^{er} et la 15^{ème} itération quand tous les 11 paramètres du modèle furent ajustés simultanément. Les paramètres trouvés furent : $a_1 = 5,5$ mm, $a_2 = 63$ mm, $a_3 = 98$ mm, $\epsilon_1 = 0,2$, $\epsilon_2 = 0,3$.

En accord avec les règles de décision du Tableau 1, cet objet fut classé comme une lettre plate de *longueur* = 196 mm, *largeur* = 126 mm et *épaisseur* = 11 mm.

Figure 5 : Reconnaissance de forme d'une lettre irrégulière comme une banane. Une banane est certainement un courrier tout à fait inhabituel, cependant un article publié dans le *Hartford Courant* le 10 Octobre 1987 a rapporté qu'une banane mûre comportant des timbres et une adresse était arrivée par le courrier régulier pour un patient dans un hôpital de New Haven. En haut de la figure se trouve l'image d'origine. En dessous se trouvent l'estimation initiale du modèle (E) et le modèle reconnu (N-30) après 30 itérations, sans utiliser aucune déformation. Le niveau de correspondance de ce modèle (N-30) est très faible. Une indication de l'objet est un colis pos-

tal irrégulier. Le modèle trouvé en utilisant la déformation de courbure atteint un meilleur niveau de correspondance—on montre la 30^{ème} itération (30) de la séquence de reconnaissance de modèle quand un total de 13 paramètres de modèle, incluant les deux paramètres de courbure, furent ajustés simultanément. Les paramètres trouvés furent $a_1 = 13$ mm, $a_2 = 19$ mm, $a_3 = 87$ mm, $\epsilon_1 = 0,6$, $\epsilon_2 = 0,7$, *rayon de courbure* = 85 mm.

Figure 6 : Influence de la finesse des grilles sur les modèles trouvés. A gauche, de haut en bas, on trouve l'image d'origine (A) et les versions de cette image sur des grilles à mailles plus lâches, obtenues en prenant un point sur deux (B), sur quatre (C) et sur huit (D) dans les axes x et y .

A droite on trouve les modèles reconnus correspondant aux images sous-échantillonnées, montrés en comparaison de l'image d'origine A.

Figure 7 : Le reste durant la reconnaissance des modèles A, B, C et D sur la figure 6, fonction du temps CPU sur un calculateur VAX 785. Les brisures de la fonction sont dues à l'addition de bruit de poisson après chaque itération ce qui permettait d'éviter le minima local peu profond. Quand la fonction de correspondance atteignait un plateau, le modèle correspondant ne s'améliorait pas du tout. La ligne en pointillé (MR) montre le reste pour une technique de correspondance par multi-résolution, au départ de la reconnaissance de modèle avec la grille d'image la plus lâche et lors de la commutation sur une grille plus fine quand la fonction de correspondance ne s'améliorait plus du tout.

ERMITTLUNG DER FORM VON POSTSENDUNGEN AUFGRUND VON MASSBILDERN MITTELS DREIDIMENSIONALER VERFORMBARER MODELLE

Franz Solina, Ruzena Bajcsy

Bei der automatischen Verteilung aller möglichen Sendungsarten sind die unterschiedlichen Formen und Größen der Sendungen zu berücksichtigen. Hier wird ein System vorgestellt, das die Unterscheidung der Sendungen nach ihrer Form mit Hilfe von Bildern ermöglicht, die die Ausmaße der Sendungen wiedergeben. Hierbei werden durch Verkleinerung auf der Grundlage der Methode der kleinsten Quadrate anhand einer Modellpaßfunktion volumetrische Modelle der einzelnen Sendungen mit Hilfe der charakteristischen Parameter Lage, Ausrichtung, Größe und Form erstellt. Bei den Modellen handelt es sich um Superquadriken mit globalen Verformungen. Die ermittelten Parameter sind je nach der für die Manipulation des Sendungsstromes verwendeten Methode für verschiedene Unterscheidungssysteme verwendbar. Hier wird ein Unterscheidungssystem dargestellt, mit dessen Hilfe Sendungen auf die Kategorien Flach-, Rollen-, Schachtelsendungen und Sendungen mit unregelmäßigen Formen verteilt werden.

RECUPERO DELLA FORMA DEGLI OGGETTI POSTALI DA IMMAGINI DI RIFERIMENTO CHE UTILIZZANO MODELLI TRIDIMENSIONALI DEFORMABILI

Franz Solina, Ruzena Bajcsy

Lo smistamento automatico dei vari tipi di corrispondenza deve tener conto delle diverse forme e dimensioni degli oggetti postali. Viene proposto un sistema per la classificazione degli oggetti postali secondo la loro forma facendo uso di immagini di riferimento (range images). I modelli volumetrici di singoli oggetti postali caratterizzati dai parametri relativi a posizione, orientamento, dimensione e forma vengono costruiti con il metodo dei minimi quadrati. I modelli usati sono superquadrici con deformazioni globali. I parametri costruiti possono servire per differenti schemi di classificazione che dipendono dal particolare metodo usato per la manipolazione del flusso della corrispondenza. Viene presentato uno schema di classificazione che divide gli oggetti postali in oggetti piani, cilindri, scatole e oggetti irregolari.

立体可変モデルを使用して、範囲イメージから 郵便物の形状を再生する方法

F. ソリナ、R. Bajcsy

全種類の郵便物を自動的に区分する場合には、郵便物の形状及び大きさの違いを考慮しなければならない。範囲イメージを使用して、形状により郵便物を分類するシステムを提案する。位置、方位、大きさ、形状パラメータで表される。各郵便物の容量モデルを、適切な関数モデルの最小2乗法を利用して再生する。使用されるモデルは、全変形を伴う2次超局面である。再生されるパラメータは、郵便物の流れを操作するために使われる特殊方法に従属する。種々の分類体系として活用可能である。郵便物を平型、筒型、箱型、その他に区別する分類体系が提示される。